Measuring the degree distribution of routers in the core internet

Matthieu Latapy, Elie Rotenberg, Christophe Crespelle, Fabien Tarissan

Measuring the Degree Distribution of Routers in the Core Internet

Matthieu Latapy[1] Élie Rotenberg[1,2] Christophe Crespelle[2] Fabien Tarissan[1]

1. Sorbonne Universités, UPMC Université Paris 6 and CNRS, UMR 7606, LIP6, Paris, France

2. Université Claude Bernard Lyon 1, DANTE/INRIA LIP UMR CNRS 5668, ENS de Lyon, Université de Lyon

***Abstract*—Most current models of the internet rely on knowledge of the degree distribution of its core routers, which plays a key role for simulation purposes. In practice, this distribution is usually observed directly on maps known to be partial, biased and erroneous. This raises serious concerns on the true knowledge one may have of this key property. Here, we design an original measurement approach targeting reliable estimation of the degree distribution of core routers, without resorting to any map. It consists in sampling random core routers and precisely estimate their degree thanks to probes sent from many distributed monitors. We run and assess a large-scale measurement following this approach, carefully controlling and correcting bias and errors encountered in practice. The estimate we obtain is much more reliable than previous knowledge, and it shows that the true degree distribution is very different from all current assumptions.**

I. Introduction

The internet has become a crucial infrastructure sustaining our social, economic, cultural and scientific lives at both local and worldwide scales. Despite this, due to its history, its decentralized nature and its mere complexity, our understanding of its global structure remains very limited. In particular, it is now clear that precise knowledge of its components (devices, connections, protocols, etc) is not sufficient to understand its global structure. As a consequence, much effort is nowadays devoted to measurements of the internet, aimed at shedding light on these features [1], [2], [3], [4], [5].

One of the main approaches consists in modeling the internet as a graph where nodes are ASes, routers, end-hosts, and/or other devices, and links are physical connections, AS peering, IP neighborhood, etc. One then conducts measurements based typically on traceroute, BGP and/or anti-aliasing in order to build maps of the internet [2], [6], [7], [1]. These maps are *partial* views of the corresponding graphs, and the underlying object is not always clearly defined [8]. In addition, such maps may be *biased* by the measurement procedure [9], [10], [11], [12], [13], [14]. They contain indeed much *erroneous* data, due for instance to silent routers, dynamic routing (load balancing in particular), incorrect anti-aliasing [15], [16], [17]. This means that the properties of obtained maps may differ very significantly from the properties of the true graph, in a way that is extremely difficult to assess and correct.

ISBN 978-3-901882-58-6

We explore here a completely new approach, based on the idea that one does not need a map to estimate a given property of interest. Instead, we propose to design and perform a measurement procedure targeting the estimation of a specific property. The challenge is then to ensure that the measurement succeeds in giving a reliable estimate.

We focus on the degree distribution of core routers, *i.e.* the fraction of core routers with k links for any k. The links we consider here are the *physical links* of the router, identified by its IP interfaces. We design a measurement procedure able to reliably estimate this distribution. We then develop tools needed to run it, and perform a large-scale measurement from hundreds of monitors distributed in the internet. We obtain this way an estimate of the degree distribution of routers that is much more reliable than previous knowledge, without resorting to a map at any stage.

This paper is organized as follows. First, we present our approach in Section II and explore its theoretical relevance through simulations in Section III. Then we detail the key elements of the practical implementation: the selection and assessment of a monitor set in Section IV, the sampling of random targets and the selection of relevant ones in Section V, and the derivation of an unbiased estimate from the measurement in Section VI. We finally run our practical measurement in Section VII, we present obtained results in Section VIII, and we assess them in Section VIII.

II. Our approach

Let us consider an IP address t, which we call *target*, and let us denote by $r(t)$ the node (router or end-host) to which t belongs. RFCs [18] and [19] state that when a monitor m sends an UDP packet with destination t on an unallocated port, then $r(t)$ should answer with an ICMP Destination Unreachable (Code 3/Port unreachable) packet to m. An important detail is that the source of this ICMP packet is in principle the IP address of the interface i by which $r(t)$ sent it (see Fig. 1).

Let us temporarily assume that $r(t)$ implements this feature correctly (we handle other cases below). Now consider a set M of monitors which all send such a probe towards IP address t. If for each interface i of $r(t)$ there is a monitor m in M to which $r(t)$ answers using i, then one obtains the set of *all* interfaces of $r(t)$, and so its degree. This constitutes our basic

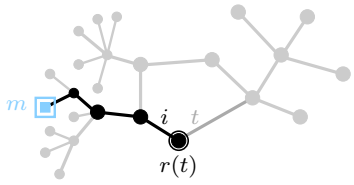


Fig. 1. Monitor m sends a UDP packet with destination address t on an unallocated port; the node $r(t)$ answers with an ICMP packet with source address i, and thus m discovers interface i of $r(t)$.

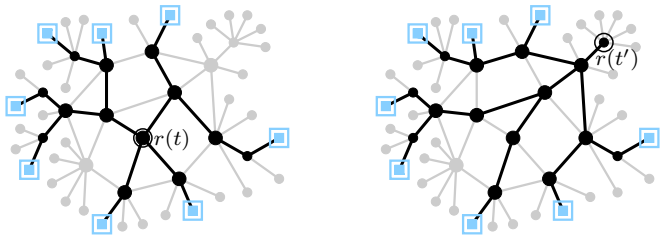


Fig. 2. Left: a set of monitors (the squared nodes) send probes towards a target IP address t and obtain the four interfaces of router $r(t)$. Right: the same monitors send probes towards another target t' but miss most interfaces of $r(t')$.

measurement primitive [1]: 1) from each monitor of a set M, we send a UDP packet to an unallocated port of target IP address t and 2) we collect the set $M(t)$ of all IP addresses used by $r(t)$ to answer to monitors in M.

Depending on the target t and on the set of monitors M this measurement primitive may succeed or fail to discover all interfaces of $r(t)$. In particular, one has to distinguish between two drastically different kinds of targets: 1) the target node $r(t)$ is in the *core* internet, see Fig. 2 (left) or 2) the target node $r(t)$ is in the *border*, see Fig. 2 (right). This distinction deserves more attention.

Given a graph, let us consider the following pruning process: iteratively remove all nodes having degree one until there remains no such nodes. We consider border nodes as being the ones removed when this process is applied to the physical internet topology. Core routers are the others. They necessarily have more than one interface linking them to another core routers, and we call such interfaces *core interfaces*. We call *border interfaces* all other interfaces, *core degree* (resp. *border degree*) of a node its number of core (resp. border) interfaces, and we call *branching points* the core routers that have at least one border interface. For instance, in Fig. 2, $r(t)$ is a core router, $r(t')$ is a border node, and the black node directly linked to $r(t')$ is at the same time a core router and a branching point.

As illustrated in Fig. 2 (right), when the target address belongs to a border node our measurement primitive misses most of its interfaces, and most likely discovers only the interface directed towards the core. This is not an issue here, as we focus on core routers, which form the key part of the network. We will see in Section V how to decide whether a target address belongs to a border node or not.

The situation regarding core interfaces of core routers is quite different. Indeed, such interfaces are not only used to communicate locally with a part of the border; in principle, they route traffic toward a non-negligible part of the internet, and one may therefore expect that a reasonably large and well distributed set M of monitors discovers them. Of course, this highly depends on the considered set of monitors and on the topology of the network. This is investigated in depth in Sections III and IV.

In summary, we expect a good enough set of monitors M to be able to discover all or almost all core interfaces of any core router, leading to an estimate of its degree in the core internet topology. Now if we consider a set T of targets sampled uniformly at random, independently from their degrees (which is discussed in Sections VI and V), then the distribution of degrees observed in T is an estimate of the degree distribution of core routers (which is more and more accurate as T grows).

Finally, our method to estimate the degree distribution of internet core routers consists in four steps:

1) obtain a large and well distributed set M of monitors,
2) build a large set T of random target addresses belonging to core routers,
3) estimate the degree of $r(t)$ for each target t in T using our measurement primitive,
4) derive from this our estimate of the degree distribution.

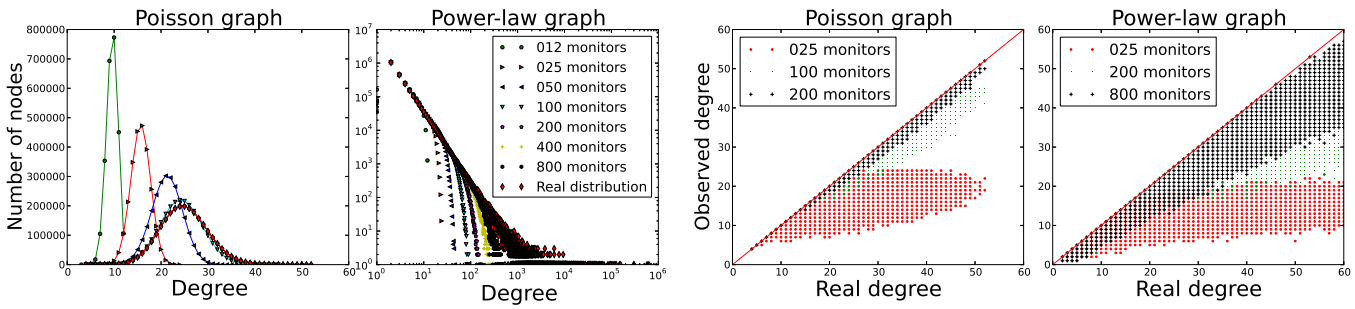
III. Proof of concept

Before putting our approach into practice, we first assess it using simulations in this section. Assuming that we are able to build appropriate sets of monitors and targets, the key questions we want to answer are: what is the risk that our estimate of a node's degree is different from its real degree, and how many monitors do we need to have an accurate estimate of the degree distribution?

To investigate this, we have conducted simulations as follows (see [20] for more details): we considered different kinds of artificial graphs to model the topology; we used as monitors random nodes with degree one (representing end-hosts); and we used *all* core targets (*i.e.* nodes in the graph obtained by iteratively removing degree-one nodes). We then assumed that each target answers to probes from each monitor using one (randomly chosen) of its interface that starts a shortest path from the target to the monitor. We used two different kinds of topologies: one with Poisson degree distribution, which is a typical homogeneous distribution, and one with a power-law degree distribution, which is a typical heterogeneous distribution. These two kinds of distributions are considered as extreme cases for what the actual degree distribution may be.

Fig. 3 shows the results of the simulations for Poisson and power-law graphs of 2.5 million nodes. Fig. 3(a) presents the degree distribution observed with respectively 12, 25, 50, 100, 200, 400 and 800 monitors. As one could expect, with 12 monitors the degree distribution is poorly estimated

[1] This is the converse of a classical *anti-aliasing* technique, aimed at identifying IP addresses belonging to a same node in a given set of IP addresses, see Section X.



(a) Observed degree distribution

(b) Scatter plot of the real degree vs. the observed degree

Fig. 3. Simulations with different number of monitors (12, 25, 50, 100, 200, 400 and 800) over graphs of $2.5 * 10^6$ nodes whose degree distribution follows either a Poisson law with average degree 25 or a power law with exponent 2.1.

in the two cases. Nevertheless, it is remarkable that, even with this poor level of quality, the nature of the distribution (*i.e.* homogeneous or heterogeneous) appears clearly. When the number of monitors increases, so does the quality of the observed degree distribution.

With 200 monitors in particular, the observed and the real distributions become visually indistinguishable in the homogeneous case (left). For the heterogeneous case (right), one can observe a cut-off for very large degrees. As we mentioned previously, this comes from the limitation of our method we identified *a priori*: the observed degree cannot exceed the number of monitors, and more generally, the estimate becomes inaccurate for targets whose degree is close to the number of monitors. On the other hand, for reasonably low-degree targets, lets say up to 20, the observed distribution and the real one are visually indistinguishable with 200 monitors.

These last statements are strengthened by the plots on Fig. 3(b) which shows the scatter plot of real degree (on the x-axis) and observed degree (on the y-axis) for all targets and for the two kinds of topologies. We can see that with 200 monitors, the estimate degree of all nodes is quite close to its real degree for the Poisson graphs, thus proving that our method performs very well on this kind of topology. As regards power-law graphs, we can see that using 200 monitors, the estimate degree of low-degree nodes is quite close to the real one. More than 95% of degree-2 nodes are correctly observed and this proportion drops to 85% when considering all nodes whose degree is lower than 10. This shows that, for this type of nodes at least, our method performs also very well on power-law graphs.

Therefore, the only limitation of our method in this theoretical setup seems to be the estimation of the degree of high-degree nodes in power-law graphs. Indeed, an intrinsic limitation of our method is that we cannot obtain a degree estimate larger than the number of monitor $|M|$. However, this limitation has to be put in perspective as Fig. 3(b) shows that, even if poorly estimated, they still cannot be confused with low-degree nodes. Whatever the number of monitors, the worst estimation (lower point on the y-axis) increases as the real degree increases. With 200 monitors for instance, the worst estimate of a node with degree higher than 1000 is 136.

In conclusion, both for Poisson graphs and power-law graphs, the nature and the shape of the degree distribution are correctly observed even with a low number of monitors. In addition, the observed distribution quickly converges to the real one when the number of monitors grows. The real degree of low-degree nodes is correctly observed (also true for high-degree nodes in the homogeneous case), and a high-degree node is never observed as a low-degree node.

These remarks will turn out to be crucial in Section VIII. However, the reader may wonder if these results still hold for graphs of different sizes and with different parameters, average degree for Poisson graphs and exponent for power-law graphs. These questions were investigated in [20], as well as the influence of some other parameters of the simulations. It turns out that the conclusions we derive here are still valid for different sizes and parameters. In particular, [20] shows that the size of the graph has very little importance, if any, for the quality of the observation with a given number of monitors. Then, the conclusion obtained by simulations on graphs of a few millions of nodes still holds for graphs of the size of the internet.

IV. Monitors

Our method relies on the use of a large set M of monitors distributed in the internet. It is crucial that this set is large enough since the accuracy of the estimation of the degrees of targets highly depends on this number (see Section III). On the other hand, having several monitors in the same location (typically having the same branching point) has limited interest: it is probable that most targets use the same interface to answer probes coming from these monitors (see Fig. 4). Assessing the quality of a given set M of monitors (regarding our measurement goals) is therefore crucial, and we propose here three different and complementary approaches to do so.

A. Colocated monitors

First notice that any monitor m may in principle be able to identify its branching point (*i.e.* the branching point between itself and core nodes, see Section II). Indeed, suppose that m iteratively sends k packets to k random IP addresses (for

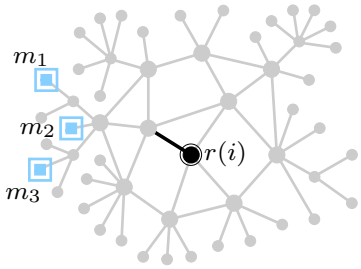


Fig. 4. Three monitors, m_1, m_2 and m_3 are actually colocated, and therefore they may observe a unique interface for any given target router $r(i)$. They are redundant regarding the quality of the measurement.

a given integer k) with increasing TTLs: the first k packets are sent with TTL 1, the k next packets with TTL 2, and so on. Thanks to the ICMP Time-Exceeded packets issued by the nodes at distance t from m (we discuss below the case of machines that do not send such packets), for each value t of the TTL m discovers a set of interfaces at distance t from m. We denote this set of interfaces by $i_t(m)$. Let us denote by $d(m)$ the smallest t such that $|i_t(m)| > 1$: $d(m)$ is the first TTL at which m discovers more than just one interface. We have by definition $|i_{d(m)}(m)| > 1$ and $|i_j(m)| = 1$ for all $j < d(m)$. Then, the (unique) interface seen by m with TTL $d(m) - 1$, *i.e.* the unique element of $i_{d(m)}(m))$ is an interface of its branching point. See for instance the case of monitor m_1 in Fig. 4, for which $d(m_1) = 3$.

Now, let us consider two monitors m and m' such that $i_{d(m)}(m) = i_{d(m')}(m')$. In other words, the first time m and m' see several interfaces they see the exact same ones. Then certainly having both m and m' in the monitor set has little interest for our measurements: m and m' enter in the core internet through very close routers (probably through the same branching point, see Fig. 4)[2]. We say that such monitors are *colocated*. The number of non-colocated monitors in M is a key value for estimating the quality of M: it basically represents the number of significantly different locations hosting monitors in M.

In the scheme we just described, we ignored machines that do not send ICMP Time-Exceeded packets. Because of them, we may erroneously decide that some monitors are colocated; this means that we under-estimate the quality of our monitor set, which has no important consequence in our context: the quality is only under-estimated. Similarly, it is possible that two monitors m and m' have different branching points but satisfy $i_{d(m)}(m) = i_{d(m')}(m')$. Again, this would make us under-estimate the quality of the monitor set and therefore we may safely ignore this. Conversely, some monitors m and m' may have different but similar sets $i_{d(m)}(m)$ and $i_{d(m')}(m')$, indicating that they are not colocated but located close from each other. It may be interesting to use this for a more subtle assessment of the level of distribution of monitors, but we leave this for further work.

[2]Notice that this does not mean that such monitors have no interest at all and should be discarded: they may lead to observation of different interfaces of the target, in particular if it implements per-destination load-balancing [17].

B. Diversity of views

In the approach above, we estimate an intrinsic quality of a monitor set M as the number of different locations hosting a monitor. A complementary view is obtained by evaluating the quality of a measurement from M towards targets in a set T. For instance, one may evaluate the quality of M as the number of distinct interfaces observed from M: $Q_0(M) = \sum_{t \in T} |M(t)|$. Clearly, if $Q_0(M') > Q_0(M)$ then M' may be considered as better than M. More subtle quality functions may be defined. In particular, it is interesting to take into account the fact that interfaces of low-degree routers are easier to observe than the ones of high-degree routers. This leads to the quality function $Q_1(M) = \sum_{t \in T} |M(t)|d(t)$ where $d(t)$ stands for the degree of target router $r(t)$. Of course we do not have the value of $d(t)$ and approximate it using the results of our measurements.

Given a quality function Q like the ones above, one may assess the impact of the addition of a new monitor m to the current monitor set, by calculating $Q(M)$ and $Q(M \cup \{m\})$. Ideally, one wants to maximize Q to collect the most accurate set of observed interfaces while keeping M as small as possible to prevent redundant measurements (which may be costly).

In practice, we will want to assess a given monitor set M, and to do so we will start from an empty monitor set and compute the expected quality improvement when monitors are added one by one, in a random order. The quality is expected to grow with the number of monitors, and then to reach a steady or almost steady regime meaning that adding more monitors would not improve the measurement significantly. Of course, if many monitors are colocated (for instance, if they are all at the same location), the quality will have precisely this behavior (as adding more monitors at the same location does not significantly improve the measurement). This is why this quality function approach is *complementary* to the colocation-based one: we will perform first the colocation and then plot the behavior of the quality function when non-colocated monitors are added, see Section IX-A.

C. Convergence of observations

Last but not least, a clear way to assess the quality of a given monitor set regarding our measurements objectives is to directly observe how the observed fraction p_k of routers of degree k converges when the number of monitors grows, for all k. Here again, we expect these fractions to converge rapidly to a steady value, which is our final estimate. This would indicate that the last monitors we added were not necessary, and thus that we obtain an accurate view. For the same reasons as above, this is complementary to colocation analysis.

V. Targets

Being able to sample a core router uniformly at random in the internet[3] would help us much, but there is no direct way

[3]*Uniformly at random* means that all possible elements are sampled with the same probability.

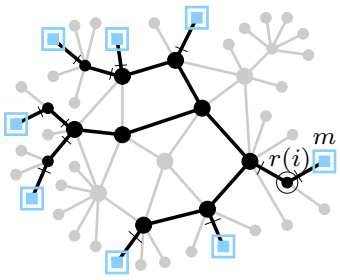


Fig. 5. If we target an interface i that belongs to a border router $r(i)$ then our measurements may see more than one interface for $r(i)$. However, only one of them does not belong to $B(M)$, as displayed in this picture: all interfaces of $B(M)$ are marked with a small dash.

to do so. Instead, it is trivial to get IP addresses uniformly at random, as they are nothing but 32 bit integers. Of course, sampling such a random integer does not necessarily give a relevant IP address with regards to our measurement needs: it may for instance belong to an end-host or a router that does not answer our probes.

In this section, we show how to sample uniformly at random an interface of an internet core router that correctly answers our probes, which we call a *correct core router*. From this sampling, which is not a *uniform* sampling of core routers themselves but only of the interfaces of some of them, we show in Section VI how to estimate the degree distribution of all internet core routers.

First notice that a core router may give incorrect answers to our probes. In particular, it may give no answer at all, or it may always answer using the same interface independently of the monitor[4]. In these cases, our measurement procedure discovers zero or one interface for the corresponding target. Instead, if the target address belongs to a correct core router, our measurements see at least two of its interfaces (as long as monitors are reasonably well distributed). Therefore, we are able to distinguish between correct core routers and other core routers.

There is no reason to assume that the degree of core routers is correlated to whether they answer correctly to our probes or not. Indeed, low-degree core routers may *a priori* misbehave as well as high-degree ones, and conversely. As a consequence, the degree distribution of correct core routers is the same as the degree distribution of all core routers. We therefore focus on correct core routers here.

Let us now consider the IP address i corresponding to a 32 bit integer sampled uniformly at random. If it belongs to a known class of reserved addresses [21], if it belongs to no machine in the internet, if it belongs to a machine that does not answer to our probes, or if it belongs to an end-host, then our measurements see only one or zero interface for it: $|M(i)| \leq 1$. As a consequence, we are able to distinguish between these cases and the one where i belongs to a correct core router.

If the target address i belongs to a border router $r(i)$, then in most cases (see Fig. 2 (right)) our measurements see only one interface. In some cases, though, we may see more than just one interface, see Fig. 5. Indeed, let us denote by $B(M)$ the set of all interfaces seen between monitors in M and the core internet in the process described in Section IV-A: with the notations of this section, $B(M) = \cup_{m \in M} \cup_{k<d(m)} i_k(m)$. By construction, all IP addresses in $B(M)$ belong to border routers, and they are all such interfaces one may observe from monitors in M, see Fig. 5. Conversely, if the target address belongs to a border router, then this router may have interfaces in $B(M)$, and these interfaces are seen from monitors in M. The key point here is that, our measurements see only one interface not in $B(M)$ for such routers. Therefore, we are able to distinguish them from correct core routers (for which we observe at least two interfaces not in $B(M)$).

In summary, we build target sets as follows. We sample random 32 bit integers and select the corresponding IP address i if and only if probes to i lead to observation of at least two interface not in $B(M)$. Such an IP address is called a *valid target*. It is sampled uniformly at random among interfaces of correct core internet routers.

VI. Bias correction

The procedure described in previous section samples uniformly at random IP addresses of interfaces of correct core routers, which we assume to be representative of all core routers. However, it does not sample uniformly at random correct core routers themselves: one has k possibilities to sample a router with k interfaces, so high-degree routers appear with probability higher than low-degree ones. More precisely, the probability to sample a router is proportional to its degree k, and so the observed fraction p'_k of routers sampled with this bias having degree k is proportional to k times the fraction p_k of routers sampled uniformly at random with degree k: $p'_k \sim k \cdot p_k$. As a consequence, we obtain:

$$p_k = \frac{p'_k}{k} \cdot \frac{1}{\sum_{i>1} \frac{p'_i}{i}}$$

where the second term is nothing but a normalization constant to ensure that $\sum_k p_k = 1$.

We may therefore use this formula to infer the true degree distribution p_k from the observed one p'_k. However, p'_k is the fraction of core routers with k *core* interfaces: our measurements see the core interfaces of core routers, not their border interfaces (see Section II). We therefore have to ensure that the target generation procedure described in previous section samples *core* interfaces (of core routers) uniformly at random. To obtain this, we discard targets that turn out to be border interfaces. We detect them as follows: they are not observed during our measurements except if they belong to $B(M)$. In other words, a target interface i of a correct core router is a border interface if and only if $i \notin M(i)$ or $i \in B(M)$.

Finally, in addition to the sampling procedure described in Section V, we discard these targets. We then get from the other targets the value of p'_k and infer the unbiased p_k using the formula above.

[4]Of course, more intricate behaviors are also possible, but they are very unlikely [16] and we ignore them here.

Notice that the sampling bias towards high-degree routers has an important benefit. Indeed, we expect high-degree routers to be relatively rare (which will be confirmed by our measurements, see Section VIII) and thus we may miss them. Uniform sampling would indeed lead to a probability p_k to sample a router with degree k, but with our biased sampling this probability is proportional to $k \cdot p_k$, and thus higher for high-degree routers. This leads to a better estimate of p_k when k is large, while for small values of k the quality of the estimate is ensured by the prevalence of low-degree routers.

VII. Measurement

We present in this section a practical measurement we conducted following our approach. We describe the whole procedure step by step, as well as the obtained dataset.

We first built an initial target set by sending (from a machine in our lab) a probe to the IP addresses corresponding to 32 bit integers sampled uniformly at random. We stopped this process when we obtained correct answers (*i.e.* ICMP Destination Unreachable (Code 3/Port unreachable)) from 3 millions such targets (we considered that no answer would arrive after 1 minute). This took approximately 10 hours.

Our initial monitor set was composed of the approximately 700 machines of the PlanetLab platform [4], which is a distributed infrastructure provided to researchers typically to conduct network measurements. Some of these potential monitors are colocated and some do not fit our requirements (they have very poor connections, for instance, or they belong to networks that filter ICMP packets). We will handle these issues below.

Given these initial target and monitor sets, we uploaded our measurement tools and the target set to each monitor and remotely asked them to send a probe to each target (in a random order to avoid situations where targets would receive many probes in a short period of time). This lasted approximately 4 hours (and so each target received at most 700 probes during this period). In order to explore the stability of our measurements, we repeated this operation three times in a row. The whole measurement (building the target set and probing each of them from each monitor three times) took less than 24 hours, with a very reasonable load for targets and monitors. At this stage, we obtained for each target its answers to the probes from all monitors (repeated three times), which we gathered onto a local machine for analysis.

Some targets and some monitors behaved incorrectly. For instance, some targets sent several answers for a unique probe. Others answered to a few monitors only, probably because of shutdowns during measurements, very low ICMP rate limiting, or other specific reasons. Conversely, some monitors received surprisingly few answers, probably due to a very poor local connections, shutdowns, or to the fact that PlanetLab machines may be overloaded (they are shared by numerous users). To avoid potential noise due to these anomalous behaviors, we first discarded targets giving multiple answers to a probe. We then observed for each monitor the number of targets that answered its probes, and conversely for each target the number of monitors that received answers from it, see Fig. 6. These plots show that most monitors received answers from most targets, as we expected. To ensure that we only keep relevant data, we discarded monitors that received answers to less than 80% of their probes, and conversely all targets that sent answers to less than 80% of probes; this represents a minority of all monitors and targets, see Table I.

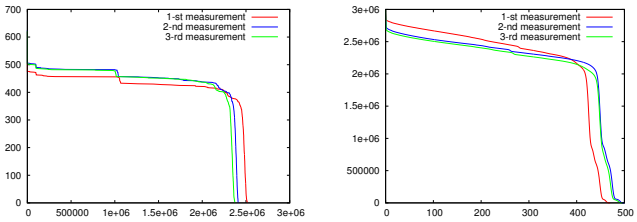


Fig. 6. Left (resp. right): for each number x on the horizontal axis, we plot the number of targets (resp. monitors) that sent (resp. received) at least x answers to our probes, for each of our three measurements.

	1-st	2-nd	3-rd
Nb running monitors	619	625	622
Nb answering targets	2849740	2734548	2699642
Nb targets answering incorrectly	10150	9842	11048
Nb monitors receiving answers from less than 80% of targets	198	183	180
Nb targets answering to less than 80% of monitors	590605	527346	544252
Nb targets t such that $t \notin M(t)$	2634226	2519320	2484483
Nb interfaces in $B(M)$	1040	1107	1097
Nb targets with only one interface not in $B(M)$	2842481	2727422	2692135
Final number of targets	5593	5623	5619

TABLE I
KEY POST-PROCESSING STEPS FOR OUR THREE MEASUREMENTS.

Following the requirements of our method, we then built the set $B(M)$ of border interface seen from our monitors and we discarded all targets t such that t is not in the set of interfaces used bt $r(t)$ to answer probes (*i.e.* $t \notin M(t)$) or t is a border interface ($t \in B(M)$), see Sections V and VI. Finally, we discarded all targets having only one interface not in $B(M)$ (which, as explained in Section V, do not belong to correct core routers).

We give the precise numbers encountered during the whole process for our three measurements in Table I.

We finally obtain for each of our three measurements approximately 5600 targets belonging to correct core routers. The key output of our measurements is the observed degree of these routers, from which we will estimate the degree distribution of internet core routers in the next section.

VIII. Results

The degree distributions observed from our three measurements, after bias correction following the formula of Section VI, are given in Table II. We plot the inverse cumulative distributions in Fig. 7.

First notice that results from each measurements are very similar, which confirms that our results are stable in this setup.

Obtained distributions show clearly that low-degree core routers are prevalent: approximately 75% of them have degree

deg	1-st	2-nd	3-rd	deg	1-st	2-nd	3-rd
2	0.74770	0.74371	0.75214	16	0.00014	0.00025	0.00024
3	0.19434	0.19838	0.19258	17	0.00023	0.00018	0.00015
4	0.02727	0.02727	0.02585	18	0.00007	0.00007	0.00007
5	0.01551	0.01588	0.01486	19	0.00007	0.00009	0.00009
6	0.00708	0.00640	0.00644	20	0.00002	0.00000	0.00002
7	0.00206	0.00224	0.00230	21	0.00008	0.00015	0.00008
8	0.00175	0.00196	0.00147	22	0.00006	0.00000	0.00004
9	0.00127	0.00131	0.00145	23	0.00000	0.00000	0.00002
10	0.00057	0.00044	0.00052	24	0.00002	0.00000	0.00002
11	0.00056	0.00052	0.00047	25	0.00000	0.00005	0.00002
12	0.00040	0.00044	0.00047	26	0.00000	0.00002	0.00002
13	0.00020	0.00023	0.00017	27	0.00002	0.00000	0.00002
14	0.00025	0.00031	0.00031	28	0.00000	0.00002	0.00000
15	0.00032	0.00009	0.00017	29	0.00002	0.00000	0.00001

TABLE II
THE DEGREE DISTRIBUTIONS OBTAINED FROM OUR THREE MEASUREMENTS (AFTER BIAS CORRECTION): FOR EACH DEGREE k, WE GIVE THE ESTIMATED FRACTION p_k OF CORE ROUTERS WITH DEGREE k.

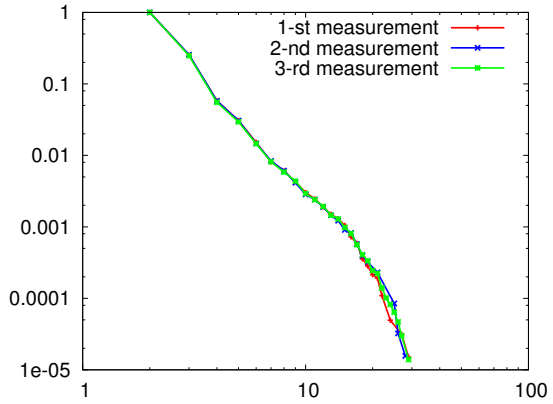


Fig. 7. Inverse cumulative degree distribution obtained from our three measurements, after bias correction: for each value x on the horizontal axis, we plot the fraction of core routers having degree higher than or equal to x (log-log scale).

2 only, and almost 20% have degree 3. This is not surprising, as we observe core interfaces only: these routers certainly have other interfaces connected to border routers and/or end-hosts. The number of interfaces they use to actually *route* traffic in the core internet, however, is very low.

Instead, some core routers have much larger degrees, and the highest one we observe is 29. We may possibly miss a few interfaces of this router but, as explained above, there is little chance that the true largest degree is much higher: we perform measurements from a much larger number of monitors and so the fact that observed degrees are bounded by this number plays no role. Of course, core routers with degree significantly higher than 29 may exist, and they probably do. There is however none in our random target set and we therefore expect them to be extremely rare (which is reinforced by the sampling bias towards high-degree routers explained at the end of Section VI).

Going further, we observe that the first values of the obtained distribution (p_k for $k < 10$) are reasonably well fitted by a power-law (a straight line in the log-log plot of Figure 7). After that, the distribution decreases less rapidly and finally it experiences a sharp decrease. The first values are the ones that our method estimates best, and so one may ask if the obtained distribution is compatible with a power-law. As highest degree may be under-estimated, this may even be in accordance with the shape of the whole obtained distribution.

In order to explore this question, we compute the range of power-law exponents compatible with the first values (the most reliable ones). We obtain $\alpha \in [3.8; 4.4]$. Beyond the actual numerical value of the exponent, this discards the usual assumption of an exponent close to 2 and this shows that *if the true degree distribution is a power-law*, it is hardly dinstinguishable from an exponential decrease in practice [22] even for a system the size of the internet.

Finally, although fully characterizing the degree distribution for large values of the degree remains to be done, our measurement shows that it significantly differs from classical assumptions: it is very heterogeneous but it experiences a much sharper decrease compatible with power-law exponents between 3.8 and 4.4.

IX. ASSESSMENT OF RESULTS

In this section, we explore two approaches to assess the quality and robustness of our results. We first study the quality of our monitor set following the methods described in Section IV. We then run simulations similar to the ones in Section III to show that our results are self-consistent.

A. Quality of the monitor set

As explained in Section IV-A, the distributed nature of our monitor set is a key feature for our measurements. We therefore ran the procedure described in this section to identify classes of colocated monitors, which provide basically redundant information. We obtained 203 different classes, each containing in average 2.11 monitors. This is consistent with the fact that each institution involved in PlanetLab often contributes with several monitors located at the same place. Examination of the DNS names of monitors belonging to a same class confirmed this: they typically match the same *.domain.tld pattern.

Once colocated monitors are identified, we investigate the diversity of views obtained from various locations, as explained in Section IV-B: we first estimate the quality of the monitor set when only one colocation class is used, then two colocation classes, etc, until all colocation classes (and thus all monitors) are used. We add colocation classes in a random order and average the obtained quality. The result is displayed in Fig. 8 (left). As expected, for both quality functions, the quality increases sharply at the beginning and rapidly converges. This indicates that adding more monitors at more locations would not improve the results much, and so that our monitor set and the number of locations hosting them are reasonable.

In order to deepen this, we examine the impact of adding more monitors at more locations on the observed fraction p_k of core routers with degree k (which is what we are interested in), as discussed in Section IV-C. We add colocation classes one by one like above and observe how p_k evolves and obtain

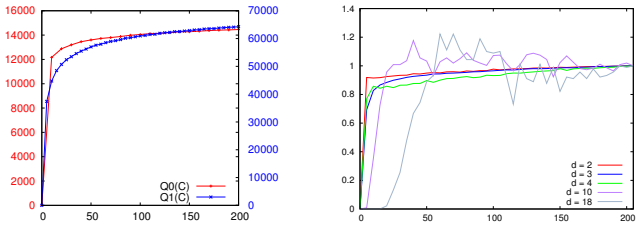


Fig. 8. Left: evolution of the quality of the monitor set when we add colocation classes. Right: convergence of the fraction of routers of degree k with the number of colocation classes.

Fig. 8. The estimates for small degrees rapidly converges, which was expected as only few monitors (and locations) are needed to observe them. Interestingly, only very few locations (approximately 10) are needed to obtain an estimate of p_k for $k < 5$ with a 80% precision. Increasing the number of monitors rapidly increases the quality of the estimate. Even for large degrees, the estimate rapidly reaches a value comparable to the final one, despite the fact that it only slowly converges after that.

Finally, this work on the monitor set shows that we have 200 significantly different locations hosting monitors, and that this is sufficient to ensure a reasonable quality for our results. It is clear however that increasing the number of monitors and the number of locations hosting them would increase both accuracy and reliability of our estimates.

B. Simulation bootstrap

We demonstrated the relevance of our approach by simulating it on artificial graphs in Section III. In the lack of a better knowledge, we used two extreme degree distributions: Poisson and power-law ones. We conduct here similar simulations but with the degree distribution obtained in Section VIII from our measurements. We expect our method to be able to observe this distribution accurately, otherwise the estimate we obtain above would make little sense.

We built 5 random graphs of 1 million nodes [5] according to each of the 3 measured distributions; these graphs represent the core internet in our simulations. For each graph, we then sampled 5 different sets of nodes at random to play the role of monitors. This leads to 75 different simulations, for which we tested sets of 12, 25, 50, 100, 200, 400 and 800 monitors. As our monitors cannot be colocated in this framework (the considered graphs have no border), the simulations most similar to our PlanetLab measurements are the one with 200 monitors.

Fig. 9 (left) displays the degree distributions observed with different sets of monitors. It shows that 200 monitors in different locations is sufficient to observe the real degree distribution, even if the fraction of high-degree nodes is less accurate than others. The plot shows that the proportion of small degree nodes is particularly well approximated: 95% of

[5]Remind that the size of the graph has little impact on the obtained results, see Section III.

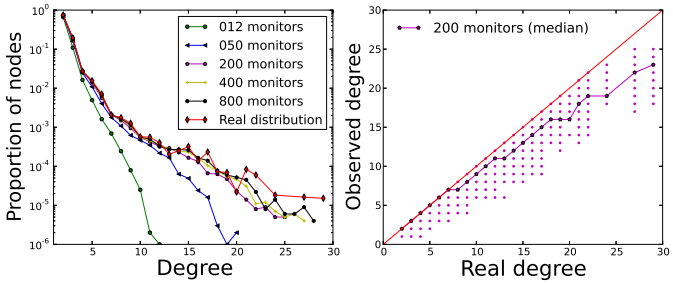


Fig. 9. Assessment by simulations. Left: the observed degree distributions with various numbers of monitors. Right: correlations between observed and real degree with 200 monitors (one dot per node and median).

all nodes with degree less than or equal to 10 are observed with their real degree with 200 monitors.

We deepen this by studying how close the observed degree of a node is to its real degree, see Fig. 9 (right). This figure confirms that our method succeeds in measuring the degree of specific nodes. In particular, the median value remains close to the real one, even for the highest degrees. Moreover, even for the highest degrees, the estimated degree is never far from the real one. For instance, 18 has been the worst estimate made for a 29-degree node; 17 for a 27-degree one and 18 for a 24-degree one. Given the fact that these are worst cases and that we cannot over-estimate a degree, such errors remain quite low.

In conclusion, simulations of our method are in accordance with our empirical measurements: the global degree distribution (which is our focus here) observed in simulations is consistent with the real one, and the estimate of the degree of specific nodes is very accurate as long as their degree is not too high. Increasing the number of monitors would provide better estimates of the fraction of high-degree nodes, without drastically changing our conclusions.

X. Related work

The physical and IP-level internet topologies are extensively studied since the seminal papers of Pansiot *et al.* [23] and Faloutsos *et al.* [24]. The most classical approach consists in building maps from traceroute-like measurements. However, several studies have shown that obtained maps are intrinsically biased [10], [11], [9], [25], [12], [13], [14], [8], and even that traceroute outputs are unreliable [17], [26], [8]. The hope that increasing the size and quality of maps would overcome these issues has led to much effort, but the situation remains far from satisfactory [9], [27], [14].

Conducting precise measurements of the degree of random nodes to obtain a reliable estimate of the degree distribution was first suggested in [10]. We explored the possibility to do so at IP level in [28] but we only partly succeeded and we conducted thorough simulations in [20]. Property-driven network measurement are also developed in other contexts, in particular Online Social Networks (OSNs) and P2P overlays.

Our work is also closely related to alias resolution (which plays a key role in the building of maps): while we seek all

(unknown) interfaces of a given router identified by one of its interfaces, alias resolution aims at identifying in a given set of interfaces the ones that belong to a same router [29], [30], [15], [16]. Probes similar to ours are used in this context, in particular by the *iffinder* tool [31], as well as other techniques. Our use of such probes was clearly inspired by these works.

Finally, important efforts are devoted to the deployment of large and distributed measurements infrastructures, which are crucial for this field of research [1], [2], [3], [4], [5]. Some of them distribute monitoring capabilities at a huge scale (typically onto thousands of end-hosts) and so are particularly promising for us [5], [2].

XI. Conclusion and discussion

In this work, we have obtained an estimate of the degree distribution of internet core routers in a rigorous way, which makes it much more reliable than previous estimates obtained from maps. To do so, we focused on the measurement of this property rather than the collection of a large (but still partial, biased and erroneous) map of the whole internet. This made it possible to design, implement and run a measurement grounded on reasonable and well identified assumptions.

Our method also has the advantage that various assessments of its results are possible. Here, in addition to the repeated measurements, we assessed the results using variations of the monitor set and simulations. Exploring other assessment approaches would increase their reliability. In particular, one may run various anti-aliasing techniques on the results of our measurements in order to confirm that the different interfaces we discover for a given target do belong to a same router. One may also run our measurements on targets for which the true degree is known, thus providing ground truth assessment.

In another direction, one may of course use larger sets of targets in order to improve the accuracy of our estimate, in particular regarding high-degree nodes. As the measurements we presented took only 4 hours, doing so seems easy. Using more and better distributed monitors would be another important improvement. Up to our knowledge, the most promising infrastructures for doing so are DIMES and RIPE Atlas [2], [5]: they already provide thousands of well distributed monitors which fit our measurement requirements.

Finally, let us notice that our measurement method is very fast and induces only a small load both on monitors and targets. This is an important feature, which makes it possible to avoid bias due to dynamics during the measurement It also opens the way to studies of the dynamics of the degree distribution at an unprecedented time scale. Going further, one may even observe the time evolution of router interfaces and use this for better modeling of the internet and its dynamics.

Acknowledgements. This work is partly supported by the European Commission EULER project (FP7 FIRE grant 258307) and by the *Agence Nationale de la Recherche* DynGraph grant ANR-10-JCJC-0202.

References

[1] CAIDA, "Caida, macroscopic topology measurement projects," 2010, http://www.caida.org/projects/macroscopic/.
[2] Y. Shavitt and E. Shir, "Dimes: let the internet measure itself," *Computer Communication Review*, vol. 35, no. 5, 2005.
[3] H. V. Madhyastha, T. Isdal, M. Piatek, C. Dixon, T. Anderson, A. Krishnamurthy, and A. Venkataramani, "iplane: An information plane for distributed services," in *In OSDI 2006*, 2005.
[4] P. Consortium, "Planetlab: An open platform for developing, deploying and accessing planetary-scale services," 2009, .
[5] RIPE-NCC, "Ripe atlas," .
[6] A. Broido and k. claffy, "Analysis of RouteViews BGP data: policy atoms," in *NRDM*, 2001.
[7] B. Zhang, R. Liu, D. Massey, and L. Zhang, "Collecting the internet as-level topology," *ACM SIGCOMM Computer Communication Review*, vol. 35, no. 1, 2005.
[8] M. Roughan, W. Willinger, O. Maennel, D. Perouli, and R. Bush, "10 lessons from 10 years of measuring and modeling the internet's autonomous systems," *JSAC*, vol. 29, no. 9, 2011.
[9] W. Willinger, D. Alderson, and J. C. Doyle, "Mathematics and the internet: A source of enormous confusion and great potential," *Notices of the AMS*, vol. 56, no. 5, May 2009.
[10] A. Lakhina, J. W. Byers, M. Crovella, and P. Xie, "Sampling biases in ip topology measurements," in *INFOCOM*, 2003.
[11] D. Achlioptas, A. Clauset, D. Kempe, and C. Moore, "On the bias of traceroute sampling: or, power-law degree distributions in regular graphs," *J. ACM*, vol. 56, no. 4, 2009.
[12] L. Dall'Asta, J. I. Alvarez-Hamelin, A. Barrat, A. Vázquez, and A. Vespignani, "Exploring networks with traceroute-like probes: Theory and simulations," *Theor. Comput. Sci.*, vol. 355, no. 1, 2006.
[13] J.-L. Guillaume, M. Latapy, and D. Magoni, "Relevance of massively distributed explorations of the internet topology: Qualitative results," *Computer Networks*, vol. 50, no. 16, 2006.
[14] M. Latapy and C. Magnien, "Complex network measurements: Estimating the relevance of observed properties," in *INFOCOM*, 2008.
[15] M. Gunes and K. Sarac, "Importance of ip alias resolution in sampling internet topologies," in *IEEE Global Internet Symposium*, 2007.
[16] K. Keys, "Internet-scale ip alias resolution techniques," *ACM SIGCOMM Computer Communication Review*, vol. 40, no. 1, 2010.
[17] F. Viger, B. Augustin, X. Cuvellier, C. Magnien, M. Latapy, T. Friedman, and R. Teixeira, "Detection, understanding, and prevention of traceroute measurement artifacts," *Computer Networks*, vol. 52, no. 5, 2008.
[18] R. Braden, "Requirements for Internet Hosts - Communication Layers," IETF, 1989.
[19] F. Baker, "Requirements for IP Version 4 Routers," IETF, 1995.
[20] C. Crespelle and F. Tarissan, "Evaluation of a new method for measuring the internet degree distribution: Simulation results," *Computer Communications*, vol. 34, no. 5, 2011.
[21] E. Gerich, "Guidelines for Management of IP Address Space," RFC 1466 (Informational), IETF, 1993.
[22] R. Perline, "Strong, weak and false inverse power laws," *Statistical Science*, vol. 20, no. 1, 2005.
[23] J.-J. Pansiot and D. Grad, "On routes and multicast trees in the internet," *SIGCOMM Comput. Commun. Rev.*, vol. 28, no. 1, 1998.
[24] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the internet topology," in *SIGCOMM*, 1999.
[25] P. Merindol, B. Donnet, O. Bonaventure, and J.-J. Pansiot, "On the impact of layer-2 on node degree distribution," in *IMC*, 2010.
[26] B. Donnet, M. Luckie, P. Mrindol, and J. Pansiot, "Revealing MPLS tunnels obscured from traceroute," *ACM SIGCOMM Computer Communication Review (CCR)*, vol. 42, no. 2, 2012.
[27] P. Barford, A. Bestavros, J. W. Byers, and M. Crovella, "On the marginal utility of network topology measurements," in *IMW*, 2001.
[28] C. Crespelle, M. Latapy, and É. Rotenberg, "Rigorous measurement of ip-level neighborhood of internet core routers," in *NetSciCom*, 2010.
[29] R. Govindan and H. Tangmunarunkit, "Heuristics for internet map discovery," in *INFOCOM*, 2000.
[30] M. H. Gunes and K. Saraç, "Resolving ip aliases in building traceroute-based internet maps," *IEEE/ACM Trans. Netw.*, vol. 17, no. 6, 2009.
[31] B. Huffaker, D. Plummer, D. Moore, and K. Claffy, "Topology discovery by active probing," in *SAINT*, 2002.